

APPLICATION FOR UNITED STATES PATENT

in the name of

Kenneth E. Feuerman

of

Adobe Systems Incorporated

for

SELF-DESCRIBING FORMS

Fish & Richardson P.C.
500 Arguello Street, Ste. 500
Redwood City, CA 94063
Tel.: (650) 839-5070
Fax: (650) 839-5071

ATTORNEY DOCKET:
07844-612001

DATE OF DEPOSIT: November 3, 2003

EXPRESS MAIL NO.: EV 321 383 533 US

SELF-DESCRIBING FORMS

BACKGROUND OF INVENTION

This invention relates to techniques for generating and processing self-describing forms. Form processing refers to the process of extracting data from a form, such as the
5 extraction of handwritten or machine printed data from a paper-based form or the extraction of audio data from an audio-based form. For example, sales orders, credit card applications, enrollment questionnaires and surveys can all require the insertion of data onto a printed form by a user, either by handwriting or using a machine, such as a typewriter. Historically, extracting user data from a form required a human operator to read the form and manually
10 key the data into a storage system such as a database - a labor-intensive and therefore expensive and time consuming task.

With the advent of automated form processing technology, including the use of optical character recognition (OCR) and intelligent character recognition (ICR), the task has become more efficient, reducing the need for human operators. A paper-based form that
15 includes form data, that is, the information printed onto the form itself (*e.g.*, the word "Address"), and user data, that is, the information added to complete the form by a user (*e.g.*, the user's address), can be used to create an image file of the completed form. For example, the paper-based form can be image scanned to create a PDF or TIFF file. A program receives the image file as input, locates the user data, and translates the images forming the user data
20 into character codes, for example, ASCII, and may output a text file. The program can be an OCR program, which is typically used to recognize machine-printed characters, an ICR program, which is typically used to recognize handwritten characters, or a program that can perform both OCR and ICR. Hereinafter, the term "OCR/ICR program" shall be used to refer to a program that can perform either OCR, ICR or both. The OCR and ICR processes
25 typically involve complex image processing algorithms and may require manual proof reading to correct inaccuracies.

In order to distinguish between forms data and user data, information can be provided to the OCR/ICR program that identifies locations on the form where user data is expected to be found, typically referred to as zoning information. Additional information can be
30 provided, that identifies certain aspects of the user data expected to be found at a particular

location. For example, with respect to a form field requesting the user's social security number, information can be provided to the OCR/ICR program specifying that a numerical value is expected. When performing character recognition, the OCR/ICR program will therefore not mistake, for example, the number "1" with the letter "l".

5 One conventional method of making zoning and other such information accessible to an OCR/ICR program is to maintain a catalog of information related to a set of forms, which is accessible by the OCR/ICR program, for example, via a networked database. In order to use the catalog, the OCR/ICR program first identifies the form, so that the corresponding zoning information can be retrieved. A form identifier can be encoded onto the form, for
10 example, using a two-dimensional (2D) graphical symbol, such as a 2D barcode. The OCR/ICR program reads the barcode, learns the identity of the form, and looks up the corresponding zoning information in a catalog accessible by the OCR/ICR program. Alternatively, a barcode can encode a URL address, which the OCR/ICR program can use to retrieve the corresponding zoning information from a remote location, for example from the
15 location specified by the URL and using an Internet connection. The zoning information can then be used to facilitate the processing of the form, as described above.

SUMMARY

 The present invention provides methods and apparatus, including computer program products, for creating and reading forms including one or more data fields. In general, in
20 one aspect, the invention features generating a form having one or more data fields, including defining zoning information identifying a location of the one or more data fields of the form and defining structural information about the one or more data fields. The zoning and structural information is encoded according to a symbology defined by rules for encoding information in a medium in which the form will be presented. The encoded zoning and
25 structural information is incorporated in a representation of the form to be presented in the medium.

 In general, in another aspect, the invention features creating a form having one or more data fields, including generating a form definition defining the form. The form definition includes zoning information describing a location of the one or more data fields.
30 The zoning information is encoded according to a symbology defined by rules for encoding information in a medium in which the form will be presented. The encoded zoning

information is incorporated in a representation of the form to be presented in the medium. The data entered on the form by a user can be extracted from the representation based on the encoded zoning information, without access to a source of zoning information external to the form.

5 In general, in another aspect, the invention features creating a form having one or more data fields, including generating a form definition defining the form. The form definition includes an XML representation of zoning information describing a location of the one or more data fields and structural information about the one or more data fields. The XML representation of the zoning and structural information is encoded according to a two-
10 dimensional symbology defined by rules for encoding information in a visual medium in which the form will be presented. The encoded zoning and structural information is incorporated in a visual representation of the form. The data entered on the form by a user can be extracted from the representation based on the encoded zoning and structural information, without access to a source of zoning and structural information external to the
15 form.

 Implementations can include one or more of the following. The medium can be a visual medium (*e.g.*, paper) and the zoning and structural information can be encoded in a graphical symbol. A graphical symbol can be a two-dimensional symbol, for example, a two-dimensional barcode or a DataGlyph®. The medium can be an audio medium and the
20 zoning and structural information can be encoded in an audio signal. The zoning and structural information can be represented in XML and the XML representation can be encoded according to the symbology.

 Where the medium is a visual medium, the zoning information can include two-dimensional coordinates specifying a location of each of the data fields and corresponding
25 measurements in two dimensions of each of the data fields. Where the medium is an audio medium, the zoning information can include a temporal location of each of the data fields in an audio recording and temporal dimensions of each of the data fields. The structural information can include a name for each of the data fields, and/or can include a description of user data expected to filled in each of the one or more data fields (*e.g.*, numeric or alpha).
30 The data entered on the form by a user can be extracted from the representation based on the encoded zoning and structural information, without access to a source of zoning or structural information external to the form.

In general, in another aspect, the invention features receiving an electronic representation of a form including user data associated with one or more data fields. The form incorporates zoning information describing a location of the one or more data fields, and structural information about the one or more data fields. The zoning and structural information are encoded according to a symbology defined by rules for encoding information in a medium in which the form is presented to a user. The zoning and structural information is decoded, and the user data is extracted from the electronic representation of the form using the decoded zoning and structural information, without access to a source of zoning or structural information external to the electronic representation of the form.

Implementations of the invention can include one or more of the following. The medium can be a visual medium (*e.g.*, paper) and the electronic representation of the form can be a PDF file or a TIFF file. The medium can be an audio medium and the electronic representation of the form can be a digital audio file. Where the medium is a visual medium, the encoded zoning and structural information can be a graphical symbol, such as a two-dimensional symbol (*e.g.*, a two-dimensional barcode or DataGlyph). Where the medium is an audio medium, the encoded zoning and structural information can be an audio signal. The zoning and structural information can be represented in XML.

The invention can be implemented to realize one or more of the following advantages. Self-describing forms that incorporate encoded zoning and structural information in a representation of the form can be processed by an OCR/ICR program independent of zoning and structural information from a source external to the form. That is, the zoning and structural information describing the form is accessible to the OCR/ICR program from the form itself, and without requiring access to external zoning and structural information accessible, for example, from a forms catalog or website. There is no need to issue a form identification number (ID), register the ID in a catalog, maintain the catalog up-to-date and imprint the ID on the form. Delays associated with entering the information into a separate catalog or database, before the form can be processed by an OCR/ICR program, are eliminated. Additionally, because the OCR/ICR program does not need to access an external catalog or database, a machine executing the OCR/ICR program does not have to be connected, via the Internet or otherwise, to a remote source including zoning and structural information.

The zoning and structural information associated with a form can be changed and the updated information can be encoded on any subsequently generated forms. Because the encoded zoning and structural information is incorporated in the form, and therefore always consistent with the particular version of the form, there is no chance that an inconsistent version of the zoning and structural information will be used to process a form. Additionally, because the life of a specific version of a form may not be known, the requirement of maintaining a potentially large collection of form identifiers and corresponding zoning and structural information for an indeterminate amount of time is avoided.

The details of one or more embodiments of the invention are set forth in the accompanying drawings and the description below. Other features and advantages of the invention will be apparent from the description, the drawings, and the claims.

DESCRIPTION OF DRAWINGS

FIG. 1 shows a paper-based form.

FIG. 2 is flowchart showing a process for creating a form.

FIG. 3 is an XML representation of zoning and structural information.

FIG. 4 is a flowchart showing a process for processing a form.

Like reference symbols in the various drawings indicate like elements.

DETAILED DESCRIPTION

A form for collecting user data is created, including one or more data fields where a user filling in the form is expected to enter the user data. An author of the form defines zoning information identifying locations within the form of the one or more data fields, and therefore locations where user data can be expected to be found by an OCR/ICR program extracting the user data from the form. Optionally, a form author can further specify structural information that can describe the form, the data fields and/or relationships between the data fields (other than the location of the data fields, which is specifically referred to herein as zoning information). The zoning information, and optionally the structural information, is encoded according to a symbology that is defined by rules for encoding zoning and structural information in a medium in which the form will be presented to a user. The encoded zoning and structural information (*i.e.*, an encoded representation of that information) is incorporated in a representation of the form to be presented in the medium.

The user data entered by a user can be extracted from the representation based on the encoded zoning and structural information. In particular implementations, the use of encoded zoning and structural information makes it possible to extract the user data without access to a source of zoning or structural information external to the form.

5 In one implementation, the form can be presented to a user on a visual medium, for example, paper. FIG. 1 shows a paper-based form 100 including data fields 115, 120, and associated form data that textually identifies or describes information to be entered in the fields, such as the field names "Employee Name" 105 and "Social Security Number" 110. Fields 115, 120 provide locations for entering user data, such as the name 125 and the social
10 security number 130 of a user. The zoning and structural information is encoded in a graphical symbol 135 incorporated on the face of the form 100.

 In the case of a paper-based form, exemplary zoning information can include the width and height of a rectangular field where user data is expected, and coordinates in x and y planes indicating the position of, for example, the upper-left corner of the field. An
15 OCR/ICR program thereby knows where on a form to perform character recognition and does not perform unnecessary character recognition on the form data itself.

 Structural information is information describing the form, the data fields and/or the relationships between the data fields. For example, structural information can include a description of the type of user data expected to be entered in a data field to facilitate
20 character recognition, such as "alpha" or "numeric". Structural information can include a name of a field, so that the user data extracted from the field can be associated with the field name in an OCR/ICR program's output. Structural information can include the number of data fields in a form, or relationships between the fields, such as the order in which the fields appear in a visual representation of the form, or a grouping of fields to be treated as a logical
25 unit (*e.g.*, a grouping including fields named "street address", "city", "state" and "zip code").

 Where the encoded representation of the zoning and structural information is to be incorporated in a visual representation of the form, the encoded representation can be a graphical symbol. The graphical symbol encoding zoning and structural information can be any computer-generated glyph, character, token, emblem or other graphical mark that can be
30 used to encode information in a format that can be captured and decoded by an image capture device, such as a scanner or CCD (charge-coupled device, *e.g.* a digital camera), and/or an OCR/ICR program or device. The OCR/ICR program can be a standalone application, or a

component (*e.g.*, a plug-in) of a forms processing program that will be used to process the form.

In one implementation, the graphical symbol is a two-dimensional symbol, such as a stacked or matrix type 2D barcode. For example, the graphical symbol 135 shown in FIG. 1 is a stacked 2D PDF417 barcode. Traditional one-dimensional barcodes are vertically redundant, repeating the same information vertically. 2D symbols encode data vertically as well as horizontally, increasing the density of information that can be included in a barcode of a given size. A stacked barcode consists of several thin horizontal slices of regular one-dimensional barcodes stacked on top of each other, forming an array that is scanned vertically as well as horizontally. Stacked barcodes can be read with a document scanner or a CCD (*e.g.*, digital camera). Matrix barcodes encode information using fixed-width light and dark cells and are read with a CCD.

Other 2D symbologies that can be used to provide graphical symbols 135 include "DataGlyphs[®]" developed by the Palo Alto Research Center (PARC), a subsidiary of Xerox Corporation, in Palo Alto, California. A DataGlyph is a pattern of small "v"s and "r"s encoding binary data. DataGlyphs are designed to blend into an image or graphic in which they are incorporated, and can form background shapes, for example, logos, or tints behind text or graphics. DataGlyphs can be aesthetically pleasing and less obtrusive on the face of a form than a dedicated symbol, such as a barcode. A DataGlyph can be read using a document scanner or CCD (*e.g.*, digital camera).

A OCR/ICR program decodes the graphical symbol to retrieve the zoning and structural information, and uses the zoning and structural information to extract the user data 125, 130. No access to a data store, or any other source of information housing zoning or structural information that is external to the form, is required to retrieve the zoning and structural information. The form can be processed independent of any such external data store, and there is no need for a machine executing the OCR/ICR program to have network or Internet access to an external data store, nor is there a need to maintain such an external data store of zoning and structural information corresponding to a set of forms, potentially including different zoning and structural information for different versions of the same form.

FIG. 2 is a flowchart showing a method 200 for generating a form having zoning and structural information encoded in a graphical symbol incorporated on the face of the form, as shown in FIG. 1. The method can be implemented in a forms authoring program such as

Adobe® Forms Designer, available from Adobe Systems Incorporated of San Jose, California. The forms authoring program is used to create a form. The author specifies a form definition, that is, defines a plurality of fields for the form and specifies form data associated with each field (*e.g.*, Employee Name 105) and a location of the field (Step 205).
5 Additionally, the author can define structural information about the form, for example, the number of fields in the form, and structural information specific to a field. For example, the author can specify that the Employee Name field 115 has a type “alpha”, and the Social Security Number field 120 has a type “numeric”, where the type of user data expected (*i.e.*, alpha or numeric) is structural information.

10 The forms authoring program generates a description of the zoning information and structural information in a suitable format. In one implementation, the zoning and structural information can be represented in XML. FIG. 3 shows an example of an XML representation of zoning and structural information 300 corresponding to the form shown in FIG. 1. The form’s author defined a name, location and data type for each field. For
15 example, the Employee Name field 115 is represented in XML by the data string 302, including a name 305, specified by the author as “EMP_NAME”, a location specified as x and y coordinates with explicit measurement units (*e.g.*, millimeters (mm)) 310, a field size specified as width (w) and height (h) 315 (also with explicit units), and a type 320, specified as “alpha”.

20 The zoning information, that is, the x and y coordinates and the width and height of the field, can be used by an OCR/ICR program to locate user data corresponding to the Employee Name field. The field name, EMP_NAME 305 can be included by the OCR/ICR program in an XML string output by the program in association with the user data extracted from the associated location. The type (*e.g.*, alpha) can be used by the OCR/ICR program to
25 facilitate character recognition, for example, to distinguish between the number “1” and the letter “l”. An XFA (XML Forms Architecture) specification can be defined to specify a format for zoning and structural information, for example, using parts of existing specifications, such as XFA specifications for templates and datasets. The XML representation of zoning and structural information can then conform to such an XFA
30 specification.

The forms authoring program constructs an XML string incorporating the zoning and structural information, as described above, and can then optionally compress the string

using conventional text compression techniques, such as flate compression. The resulting binary data can then be encoded according to rules and algorithms of a particular symbology selected, for example, a PDF417 barcode, such as barcode 135 shown in FIG. 1 (Step 210). The result is a graphical symbol in the form of a bitmap image, and the forms authoring
5 program can prompt the author for placement of the bitmap image onto the face of the form. The graphical symbol 135 is thereby incorporated into a visual representation of the form 100 (Step 215). The form can be output as an image file, for example, a PDF file, which can be emailed to a user, or accessed by a user over a network, such as the Internet. The user can then print a paper copy of the form (complete with the graphical symbol 135) and fill in the
10 user data either by writing the data by hand, or using a machine, such as a typewriter. Alternatively, a paper copy of the form (including the graphical symbol 135) can be provided to a user in the first instance, for example, a new patient form provided to a user upon an initial visit to a doctor's office.

The selection of a particular format for the graphical symbol 135 can depend on the
15 particular application, such as the expected workflow in which the form will be used. Some graphical symbols are more robust with respect to typical workflow damage, page skewing (*e.g.*, when faxing), spillage and obliteration than others. Some graphical symbols may be more compact, taking up less space on the form, while others, such as the DataGlyph, may be less visually obtrusive or more aesthetically pleasing. A PDF417 barcode exhibits the
20 advantages of denser data representation under poor imaging circumstances, *e.g.* faxing, is an open standard and is widely used.

FIG. 4 is a flowchart showing a method 400 for extracting user data from a form that incorporates a graphical symbol encoding zoning and structural information into the form. The method can be implemented in an OCR/ICR program, such as Adobe® Capture,
25 available from Adobe Systems Incorporated of San Jose, California. A paper copy of a completed form (*i.e.*, a form in which a user has filled in one or more fields with user data) that includes a graphical symbol encoding zoning and structural information is received by a form recipient. An electronic representation of the paper form is created,, for example, an image file created by image scanning the paper copy using a document scanner to create a
30 PDF file. Alternatively, an image file of a completed form can be directly received by the recipient, for example, if the user scans and e-mails the completed form to the recipient.

An OCR/ICR program receives as input the image file of the completed form (Step 405). The OCR/ICR program decodes the graphical symbol, for example, the PDF417 2D barcode 135 on form 100, to retrieve zoning and structural information describing the form 100 (Step 410). The OCR/ICR program performs character recognition, using the zoning
5 information to locate the user data, and the structural information to facilitate translation of the user data (Step 415). For example, as described above with reference to the Employee Name field 115, the OCR/ICR program uses the x and y coordinates 310 and the width and height 315 of the field to locate the user data corresponding to the Employee Name field on the form. The OCR/ICR program uses the type, alpha 320, to facilitate character recognition.

10 The output from an OCR/ICR program can depend on the intended recipient, for example, a database application or other such application, and might be in the form of a text file or a stream of XML. Referring to the XML representation of zoning and structural information 300 shown in FIG. 3, an OCR/ICR program can output an XML string (Step 420) that looks somewhat similar to the XML representation 300, although some information
15 in the initial XML string, such as the location information (*i.e.*, x and y coordinates 310, width and height 315), may be unnecessary to subsequent processing and therefore omitted from the output stream, and additional information, such as the extracted user content (*e.g.*, “Allen B. Smith” corresponding to the Employee Name field 115) that is required for subsequent processing can be added. The XML string may be meaningless to the OCR/ICR
20 program, however, the names associated with the fields, such as EMP_NAME 305, can have meaning to the recipient program, and provide a way to identify the associated user data.

In addition to zoning and structural information, other data can be encoded in the graphical symbol, for example, instructions indicating where and how to transmit the user data extracted from a form. After decoding the instructions and extracting the user data, the
25 OCR/ICR program can export the extracted user data accordingly, for example, to a database or web server.

In the example described above, the graphical symbol encodes both zoning and structural information. However, in another implementation, the graphical symbol can encode only information identifying the location of fields where user data is expected to be
30 found. Structural information can facilitate character recognition, but is not required for an OCR/ICR program to extract user data from data fields in a form.

The implementation described above incorporated encoded zoning and structural information in a paper-based form. Other implementations are possible, including incorporating encoded zoning and structural information in an audio-based form. For example, an audio-based form can consist of audio signals recording a voice speaking a field name followed by a pause, during which a form user is expected to enter the appropriate user data by speaking (*e.g.*, stating their name). The pattern of speaking a field name followed by a pause is continued until each field name has been presented to the user, and the user has been given an opportunity to enter corresponding user data. Audio signals including the voice speaking the form data and the user's voice speaking the user data together comprise a completed form.

A audio-based forms authoring program can incorporate encoded zoning and structural information into the form, for example, in audio signals detectable and decodable by an audio recognition program used to extract the user data. The zoning information can include a temporal location and temporal dimensions for each data field in the form, *e.g.*, the time in seconds from the start of an audio recording where a data field begins and the duration of a pause provided for the user to enter user data. The structural information can be similar to the structural information provided for a paper-based form, that is, field names, types of user data expected, and the like. An audio recognition program detects and decodes the zoning and structural information, and uses the information to locate and extract the user data, in a similar manner as described above in the context of paper-based forms.

The invention and all of the functional operations described in this specification can be implemented in digital electronic circuitry, or in computer hardware, firmware, software, or in combinations of them. Apparatus of the invention can be implemented in a computer program product tangibly embodied in a machine-readable storage device for execution by a programmable processor; and method steps of the invention can be performed by a programmable processor executing a program of instructions to perform functions of the invention by operating on input data and generating output.

The invention can be implemented advantageously in one or more computer programs that are executable on a programmable system including at least one programmable processor coupled to receive data and instructions from, and to transmit data and instructions to, a data storage system, at least one input device, and at least one output device. Each computer program can be implemented in a high-level procedural or object-oriented

programming language, or in assembly or machine language if desired; and in any case, the language can be a compiled or interpreted language.

Suitable processors include, by way of example, both general and special purpose microprocessors. Generally, a processor will receive instructions and data from a read-only memory and/or a random access memory. Generally, a computer will include one or more
5 mass storage devices for storing data files; such devices include magnetic disks, such as internal hard disks and removable disks; a magneto-optical disks; and optical disks. Storage devices suitable for tangibly embodying computer program instructions and data include all forms of non-volatile memory, including by way of example semiconductor memory devices,
10 such as EPROM, EEPROM, and flash memory devices; magnetic disks such as internal hard disks and removable disks; magneto-optical disks; and CD-ROM disks. Any of the foregoing can be supplemented by, or incorporated in, ASICs (application-specific integrated circuits).

To provide for interaction with a user, the invention can be implemented on a
15 computer system having a display device such as a monitor or LCD screen for displaying information to the user and a keyboard and a pointing device such as a mouse or a trackball by which the user can provide input to the computer system. The computer system can be programmed to provide a graphical user interface through which computer programs interact with users.

20 The invention has been described in terms of particular embodiments. Other embodiments are within the scope of the following claims. For example, the steps of the invention can be performed in a different order and still achieve desirable results.

What is claimed is: